

Language of Bargaining

Mourad Heddaya
University of Chicago
mourad@uchicago.edu

Solomon Dworkin
University of Chicago
solomon.dworkin@gmail.com

Chenhao Tan
University of Chicago
chenhao@uchicago.edu

Rob Voigt
Northwestern University
robvoigt@northwestern.edu

Alexander Zentefis
Yale University
alexander.zentefis@yale.edu

Abstract

Leveraging an established exercise in negotiation education, we build a novel dataset for studying how the use of language shapes bilateral bargaining. Our dataset extends existing work in two ways: 1) we recruit participants via behavioral labs instead of crowdsourcing platforms and allow participants to negotiate through audio, enabling more naturalistic interactions; 2) we add a control setting where participants negotiate only through alternating, written numeric offers. Despite the two contrasting forms of communication, we find that the average agreed prices of the two treatments are identical. But when subjects can talk, fewer offers are exchanged, negotiations finish faster, the likelihood of reaching agreement rises, and the variance of prices at which subjects agree drops substantially. We further propose a taxonomy of speech acts in negotiation and enrich the dataset with annotated speech acts. We set up prediction tasks to predict negotiation success and find that being reactive to the arguments of the other party is advantageous over driving the negotiation.

1 Introduction

Bilateral bargaining, in the sense of a goal-oriented negotiation between two parties, is a fundamental human social behavior that takes shape in many areas of social experience. Driven by a desire to better understand this form of interaction, a rich body of work in economics and psychology has evolved to study bargaining (Rubin and Brown, 1975; Bazerman et al., 2000; Roth, 2020). However, this work has seldom paid careful attention to the use of language and its fine-grained impacts on bargaining conversations; indeed, many studies operationalize bargaining as simply the back-and-forth exchange of numerical values. Meanwhile, there is growing interest in bargaining in NLP oriented towards the goal of building dialogue systems capable of engaging in effective negotiation (Zhan

et al., 2022; Fu et al., 2023). In this work, we aim to bridge these two lines of work and develop a computational understanding of how language shapes bilateral bargaining.

To do so, building on a widely used exercise involving the bargaining over the price of a house used in negotiation education, we develop a controlled experimental environment to collect a dataset of bargaining conversations.¹ The treatment in our experiment is the manner in which subjects communicate: either through alternating, written, numeric offers (the *alternating offers* or AO condition) or unstructured, verbal communication (the *natural language* or NL condition). Furthermore, to encourage naturalistic interactions, we recruit participants via behavioral labs and allow participants to negotiate in a conversational setting using audio on Zoom instead of crowdsourcing text conversations as prior work has done (Asher et al., 2016; Lewis et al., 2017; He et al., 2018). In total, we collect a dataset with 230 alternating-offers negotiations and 178 natural language negotiations. In contrast with He et al. (2018)’s Craigslist negotiation dataset, our natural language negotiations have an average of over 4x more turns exchanged during each conversation, so our dataset represents a richer source to explore linguistic aspects of bargaining behavior than has been presented by existing work in this area.

In addition, we enrich the dataset by annotating all the conversations with a set of negotiation-specific speech acts. Inspired by prior work on rhetorical strategies in negotiations (Chang and Woo, 1994; Weigand et al., 2003; Twitchell et al., 2013), we create a simplified taxonomy of what we term *bargaining acts* and hire undergraduate research assistants to provide annotations. To the best of our knowledge, our dataset of speech acts in negotiations is an order of magnitude larger than

¹Dataset access may be requested at: <https://mheddaya.com/research/bargaining>

existing datasets.

We first provide descriptive results based on our dataset. Although the AO and NL conditions are conducted via different communication mechanisms, they reach the same average agreed prices. However, when subjects can talk, fewer offers are exchanged, negotiations finish faster, the likelihood of reaching agreement rises, and the variance of prices at which subjects agree drops substantially. These observations suggest that the use of language facilitates collaboration. We also find differences in how buyers and sellers employ bargaining acts.

Recorded and transcribed speech provides more direct access to the intuitive attitudes and behaviors of the buyers and sellers. This enables us to identify subtle types of expression that are predictive of negotiation outcomes and reveal underlying dynamics of negotiation. Other findings corroborate conclusions from Lee and Ames (2017), who distinguish the effectiveness of negotiators' different expressions of the same rationale.

We set up prediction tasks to predict the outcome of a negotiation based on features of the conversation and analyze the important features contributing to class differentiation. Our results show that LIWC features provide consistently strong performance and even outperform Longformer (Beltagy et al., 2020) given the beginning of a negotiation. Important features reveal that successful sellers address buyers directly with the pronoun "you" and prompt them to divulge their considerations and preferences, while successful buyers convey a sense of emotional indifference and thoughtfulness through filler words and curt acknowledgements of the sellers' arguments.

In summary, we make the following contributions:

- We build a novel dataset of bargaining and provide annotations of bargaining acts;
- We demonstrate that the ability to communicate using language facilitates cooperation.
- Our work reveals linguistic signals that are predictive of negotiation outcomes. For instance, it is advantageous to be patient and reactive to the arguments of the other party, rather than the one driving the negotiation.

2 Related Work

Negotiation is a growing area of study in computer science. Zhan et al. (2022) provide an excellent survey of research on negotiation dialogue systems.

Lewis et al. (2017) train recurrent neural networks to generate natural language dialogues in negotiations. He et al. (2018) propose a modular generative model based on dialogue acts. Our focus is on deriving computational understanding of how language shapes negotiation.

Several research disciplines have studied bilateral bargaining from different perspectives and using different tools. Economic theory has investigated the role of incomplete information (Ausubel et al., 2002) and highlighted the role of explicit communication (Crawford, 1990; Roth, 2020). Bazerman et al. (2000) and Pruitt (2013) provide an overview of the psychology literature on negotiation. However, these studies tend to overlook the *content* of the communication, with some notable exceptions (Swaab et al., 2011; Jeong et al., 2019; Lee and Ames, 2017).

The most related work to ours is Lee and Ames (2017), who study how bargaining outcomes are affected by the way a rationale is expressed. They find that expressions that hint at a constraint (e.g., "I can't pay more") are more effective at shaping a seller's views of the buyer's willingness to pay than critique rationales (e.g., "it's not worth more").

3 Dataset

The first contribution of our work is building the first transcript dataset of *spoken* natural language bargaining between lab experiment participants. Our dataset extends existing datasets in four ways:

1. Negotiation happens in spoken language, and is thus more fluid and natural, akin to real-world bargaining scenarios, such as price haggling in vendor markets, union negotiations, or diplomacy talks, while existing work is largely based on written exchanges (Asher et al., 2016; Lewis et al., 2017; He et al., 2018);
2. Our work is the first one to introduce a control condition without the use of natural language;
3. Participants are recruited through behavioral labs at universities and their incentive structure is more high-powered (i.e., bonus earnings based on outcomes and payment exceeding the typical \$12 hourly wage) than for a crowdworker on Amazon Mechanical Turk;
4. We supplement the transcripts with manual annotation of speech acts (see §4).

While contributing greatly to our understanding of negotiation, existing bargaining datasets are

somewhat limited in being based on written exchanges (He et al., 2018), often in the context of a highly structured game (Asher et al., 2016; Lewis et al., 2017).

Experiment design. We conducted a controlled experiment whose setting reflected a common life experience: the purchase or sale of a house. We adapted the setting in “Buying a House” by Sally Blount, a popular exercise from the Dispute Resolution Research Center (DRRC) of Northwestern University’s Kellogg School of Management (Blount, 2000).² We randomly paired participants and each was assigned the role of buyer or seller. In each pairing, buyer and seller negotiated a price of the house anonymously. Both buyer and seller were aware of the listing price of \$240,000 and shared the same descriptions of the house and surrounding area, along with recent sales prices of comparable homes. However, each participant was given a private valuation of the house (\$235,000 for the buyer and \$225,000 for the seller).

Participant bonus earnings depended on bargaining outcomes to incentivize subjects to engage in realistic negotiating behavior. If no agreement was reached, neither party earned bonus money. On an hourly basis, compensation seemed significant enough to influence participant behavior (i.e., at least \$40/hour was on the table per round). On average, subjects earned roughly \$23.25/hour. More details can be found in Appendix B.

Each subject participated in two bargaining rounds. In the first round, a buyer-seller pair communicated via *alternating offers* (AO) in an online chat that only accepted numeric entries. Each participant could choose to accept or counter each offer they received. In the second round, participants played the same role, either buyer or seller, but were assigned a new partner. In this round, each pair communicated in *natural language* (NL) via audio only on Zoom (videos were monitored to be turned off to avoid signals from gesture and facial expressions). The subjects were restricted from disclosing their respective private value and compensation structure and informed that doing so would result in forfeiture of their earnings.³ Our ex-

²Thanks to the DRRC for kindly granting us permission to base our bargaining setting on this negotiation exercise that teaches purely distributive (i.e., zero-sum) bargaining between two parties.

³To control for the order of the two treatments affecting the bargaining outcomes, roughly half the sessions (58% of the negotiations) first began with the round of alternating

	AO	NL
No. of Turns	29.2	42.50
No. of New Offers	17.9	6.06
No. of Repeat Offers	11.3	1.56
Duration (min)	9.5	6.5
Avg Turn Length (sec)	28.9	12.54
Prob. of Agreement (%)	90.0	97.19
Agreed Price (\$000s)	229.9	229.8
No. of Negotiations	230	178
No. of Unique Participants	460	356

Table 1: Descriptive Statistics Across Treatments; The table reports mean descriptive statistics of the house price negotiations in the Alternating Offer (AO) and Natural Language (NL) treatments.

periment is approved by the IRB at an anonymous institution.

Preprocessing. We transcribed the audio from the Zoom negotiation settings using Amazon Transcribe. Transcription produces strictly alternating seller and buyer turns, without sentence segmentation. We use the resulting transcripts for the annotation and analyses described in this paper. We trim the end of each negotiation at the point of agreement on a final price for the house, discarding any interaction that occurs subsequently. We describe in §4 the annotation procedures that allowed us to reliably identify the point of agreement.

Descriptive statistics. Table 1 provides descriptive statistics of the AO and NL treatments. Since a failed negotiation results in no bonus for both sides, most negotiations end with a successful sale. Nevertheless, the probability of agreement is roughly 7 percentage points higher under NL than AO (97.2% versus 90.0%). A two-tailed t-test with heteroskedasticity-robust standard errors shows that the difference in agreement probability is significant. Moreover, in contrast with the AO treatment, the NL treatment produces negotiations that, on average, have $\sim 1.5x$ more turns, but NL turns are over 50% shorter in duration, and NL negotiations are roughly 30% shorter in total duration and feature about 74% fewer offers.

Surprisingly, without the ability to communicate using language, buyers and sellers are less efficient in reconciling their differences. In the AO treatment, the combination of fewer turns that are each, individually, longer in duration is telling. Interlocutors are spending more time silently strategizing

offers, whereas the other half began with the round of natural language. We did not detect any ordering effects.

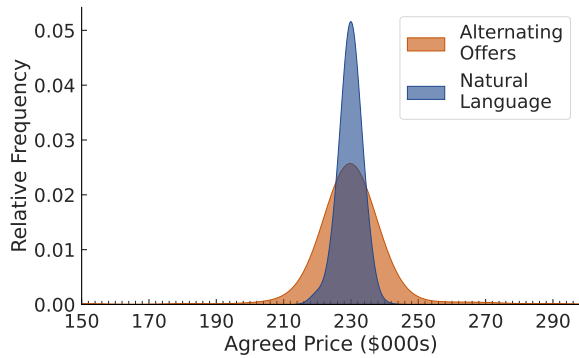


Figure 1: Gaussian kernel estimates of the distributions of agreed prices among successful negotiations.

and considering their next act. However, this time invested is not fruitful individually nor at the level of coordination, as exemplified by a lower probability of agreement and equivalent agreed prices among successful negotiations, likely due to an impoverished channel of communication.

Figure 1 shows that the distributions of agreed prices largely overlap between the two treatments, but the distribution in prices under NL is substantially narrower than under AO. Between the two treatments, the mean agreed price conditional on reaching agreement is identical (\$229.8 thousand). However, the standard deviation of agreed prices under NL is about one-third of that under AO (3.1 versus 10.4). A Fligner-Killeen (FK) (Fligner and Killeen, 1976) two-sample scale test shows that the standard deviation of the AO price distribution is statistically larger than the NL counterpart.

4 Bargaining Act Annotation

Previous researchers have recognized the inherently speech-act-like character of negotiations (Chang and Woo, 1994; Weigand et al., 2003; Twitchell et al., 2013). Many or most utterances in a bargaining context can be thought of as taking some action with reference to the negotiation. Here we propose and present a simplified ontology of negotiation-oriented speech acts (hereafter, *bargaining acts*) relevant to the present context of negotiation. Two trained undergraduate research assistants annotated all transcripts according to five *bargaining acts*: 1) new offer, 2) repeat offer, 3) push, 4) comparison, 5) allowance, and 6) end. Table 2 provides definitions and examples. Note that each turn can include multiple bargaining acts. In addition, each speech act is also annotated with a numerical offer, if applicable.

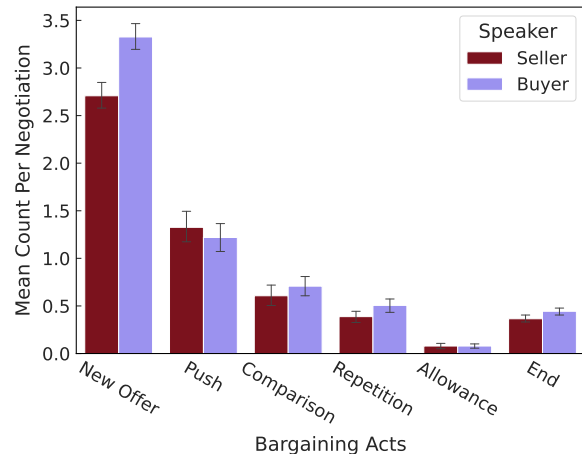


Figure 2: Distribution of Bargaining Acts. Error bars indicate standard error.

Twenty-four transcripts were annotated by both annotators to allow agreement to be calculated. Using MASI distance weighting (Passonneau, 2006), we found a Krippendorff’s alpha (Hayes and Krippendorff, 2007) of 0.72, representing a high degree of agreement for a pragmatic annotation task.

Figure 2 shows that *new offers*, *pushes*, and *comparisons* are relatively more frequent and appear more consistently in all the negotiations than *allowances* and *repeat offers*. We note in Table 1 that *repeat offers* are dramatically more common in the AO condition than the NL condition (11.3 vs. 1.56 per negotiation). With linguistic context, negotiators are less likely to engage in fundamentally uncooperative behavior by simply repeating past offers over again.

Comparing buyers to sellers, we observe that buyers make on average 1 more *new offers* per negotiation than sellers (independent sample, heteroskedasticity robust t-test, $p = 0.02$). We find no statistically significant differences between roles for the other five bargaining acts.

The bargaining act annotations allow us to describe a negotiation as a sequence of offers proposed by the buyer and seller. We compare how the frequency and pattern of numerical offers differ across 1) experimental treatments (NL vs. AO) and 2) negotiation outcomes. We characterize different properties of the negotiations as well as their *trajectories* over the course the interaction.

Figure 3 reveals three general patterns on offer trajectories. First, both AO and NL bargaining feature a similar range of new offers exchanged in the early stages of the negotiation. Early on, buyers

Bargaining Act	Definition	Example
New offer	Any numerical price, not previously mentioned, that is offered by either the buyer or seller throughout the course of the negotiation.	That’s still \$30,000 out of my budget but I would be willing to pay 210,000
Repeat Offer	Any numerical price presented that is an exact repeat of a previously presented offer; in a literal sense, these are redundant offers that were already on the table.	Yeah I understand um you still think that to 240,000 is too high right
Push	Any overt linguistic effort made by either party to bring the other party’s offer closer to theirs.	Might just be a little too low for what I have to offer here
Comparison	Evokes a difference or similarity between an aspect of the seller’s house and other external houses or considerations.	Like there’s one for 213k Which is like smaller and it’s nearby so that’s closer to our budget, we’ve seen that apartment it’s not as like it’s not as furnished and it’s kind of old and so
Allowance	Any time either party adjusts their offer price closer to the other party’s most recent offer. An allowance may be interpreted as the accompanying interaction to a successful <i>push</i> act.	I mean really like it probably should be higher than 233 but we’re willing to drop it to 233
End	End of negotiation via offer acceptance entering mutual common ground - explicitly only happens once.	Alright 228 it is

Table 2: Bargaining act annotation definitions and examples.

in both treatments present new offers as low as 170; and sellers, as high as 270. But extreme offers are more prevalent in AO than NL bargaining. Second, both the AO and NL trajectories exhibit a rhythmic pattern of low and high offers, which is familiar to real-world negotiations. The buyer’s low offer is countered by the seller’s high offer, which is then countered by the buyer’s slightly increased low offer, and so on. Third, NL bargaining takes far fewer new offers to reach agreement than AO bargaining. Figure 3b clearly demonstrates that NL negotiations converge quicker, with consecutive offers converging to within \$5K after 6 new offers. AO negotiations take over 40 new offer exchanges to reach a similar convergence.

5 Predicting Negotiation Outcomes

Finally, we set up prediction tasks to understand the relationship between the use of natural language and negotiation success. Overall, our models demonstrate performance gains over majority class in most settings. Surprisingly, Logistic Regression using bag-of-words and LIWC category features outperform the neural model. We observe differentiation between classification accuracy on seller only and buyer only speech, and highlight features that explain this difference.

5.1 Experimental Setup

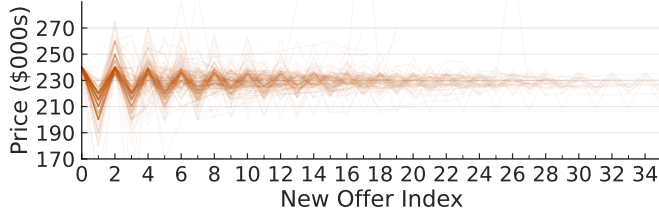
Task. We consider a binary classification task with two classes: 1) “seller win” and 2) “buyer

win”, where a negotiation is classified by whether it concluded with an agreed price greater than \$230K or less than \$230K, respectively. We focus on negotiations that end with an advantage for either the buyer or seller to better understand the dynamics that produce an asymmetric outcome. Hence, we omit the negotiations that ended with \$230K or that did not reach an agreed price. This leaves us 119 negotiations.

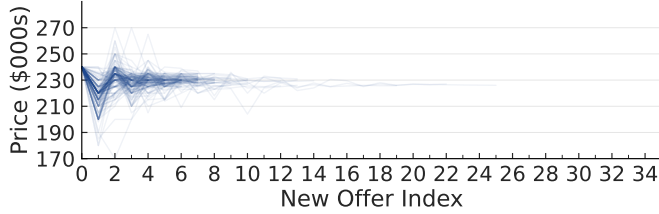
As the predictive task may become trivial if we see the entire exchange, we build 10 versions of each negotiation by incrementally adding proportions of the negotiation to the input with a step size of 10%. Thus, we obtain input/output pairs (X_k, y) for a given negotiation, where $k = \{10\%, \dots, 100\%\}$, and each k corresponds to a different prediction task; namely, whether the negotiation outcome can be predicted by the first k percentage of the interaction.

Methods. We test two setups for our task. The first is a standard linear model with logistic regression. The second is an end-to-end approach using Longformer, a transformer-based model for encoding and classifying long sequences. In particular, we use the encoder and output modules of LongformerEncoderDecoder (LED) (Beltagy et al., 2020), a variant of the original Longformer model, which can encode sequences up to 16,384 tokens in length. This exceeds the maximum input length in our dataset.

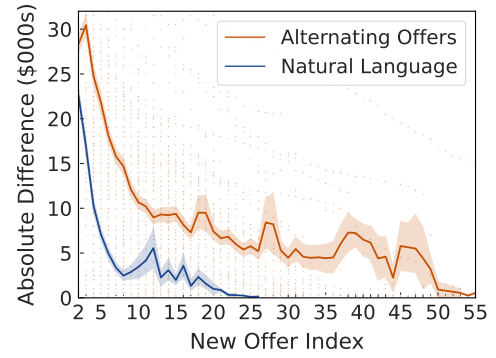
In the logistic regression experiments, we treat



(a) Alternating Offers.



(c) Natural Language.



(b) Absolute Differences in Consecutive New Offers.

Figure 3: The figure presents the trajectory of new offers in the two treatments. In 3a and 3c, each line represents a sequence of new offers exchanged between buyer and seller in a single negotiation. Only negotiations ending in agreement are included. Figure 3b presents the absolute differences in consecutive new offers under both treatments. Each dot represents an absolute difference in consecutive new offers within a single bargaining session.

the numerical offers as an oracle and consider three other feature sets: 1) Transcription texts; 2) Bargaining acts; 3) LIWC categories (Tausczik and Pennebaker, 2010).⁴ We represent each negotiation as a binary bag-of-words encoding of the features listed above. For *bargaining acts*, we construct the vocabulary based on unigrams and bigrams; for the other feature sets, we only include unigrams. We include bigrams for bargaining acts to capture local combinations of bargaining acts. To maintain a reasonable vocabulary size, we only consider unigrams from the transcribed text that occur in at least 5 negotiations (see Appendix C for total feature counts). We replace numbers mentioned in the text with a generic [NUM] token to eliminate the strongly predictive signal of new offers and focus on language instead. In experiments with LED, we add two special tokens [SELLER] and [BUYER] that we concatenate to the start of each turn depending on who is speaking. We make no other changes to the transcribed text. The input to LED is the concatenation of all the turns.

Evaluation. We use accuracy as our main evaluation metric. In all experiments, due to the relatively small size of our dataset, we use nested five-fold cross validation for both inner and outer cross validations. For logistic regression, we grid search the best ℓ_2 coefficient within $\{2^x\}$, where x ranges

⁴We also tried the union of these features, but it did not materially affect the performance.

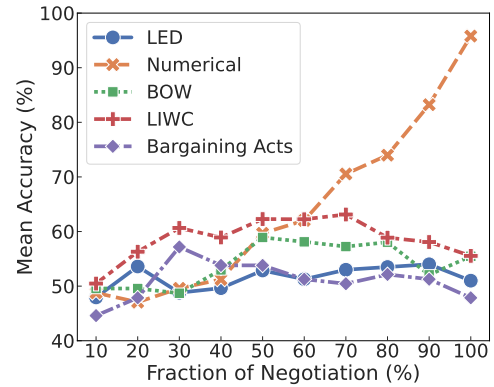


Figure 4: Overall prediction performance.

over 11 values evenly spaced between -10 and 1 . We further concatenate the speaker (‘buyer’ or ‘seller’) and the turn position within the negotiation. We treat these as hyper-parameters. We represent the position as k , where k corresponds to a fraction of the conversation, as defined earlier. For example, the word “house” spoken by the seller in the first 10% of turns in a negotiation would be tokenized as “s1-house”. In the LED experiments, we omit the inner cross validation and use a batch size of 4, the largest possible batch size given our memory constraints.⁵ We select the best performing learning rate out of $\{5e-5, 3e-4, 3e-3\}$ and early stop based on training loss convergence.

⁵We use a single Nvidia A40 GPU in our LED experiments.

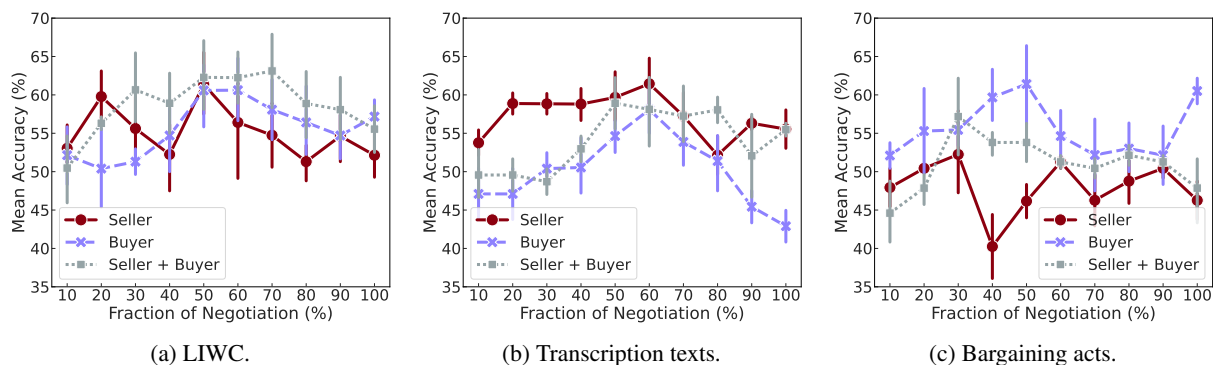


Figure 5: Buyers vs. sellers. Accuracy of Logistic Regression model across different input features, using buyer speech, seller speech, or both. Error bars indicate standard error.

5.2 Results

Predictive performance. We start by looking at the overall predictive performance. Figure 4 presents results for all models. For the oracle condition (numerical), as expected, prediction accuracy increases monotonically and steadily as the fraction of the conversation and the corresponding numerical offers in the input increases from 10% to 100% of the conversation. As the buyer and seller converge towards an agreed price, the offers made provide strong signal about the outcome.

However, this task proves much more challenging for other models where we do not include numerical offers provided by annotators. One intriguing observation is that LED consistently underperforms logistic regression. Within logistic regression, LIWC categories outperform other features and achieve 63.1% accuracy whereas text-based BOW features achieve a best score of 58.9%. Furthermore, there is no clear trend of performance growing as the fraction of negotiation increases. While bargaining actions under-perform other features overall, there is a notable jump in accuracy at fraction 30%, which we will revisit later.

Buyer vs. seller. In bilateral bargaining, an interesting question is which party drives the negotiation, and to what effect? To further understand the role of buyer vs. seller, we only consider features of buyer texts or seller texts.

Although the performance of LIWC does not vary much for buyer and seller texts (Figure 5a), Figures 5b and 5c show contrasting differences in prediction accuracy for sellers and buyers at various fractions of a negotiation. Seller transcription text achieves ~10% higher accuracy than buyer and buyer + seller at fractions 20% ($p = 0.01$), 30% ($p = 0.01$), 90% ($p =$

0.001), 100% ($p = 0.01$). Meanwhile, buyer bargaining acts outperform seller acts throughout and are particularly effective at 40% ($p = 0.008$) and 50% ($p = 0.03$) of the negotiation.

Important features. To understand in greater detail which features are more helpful for prediction, we compare the fitted logistic regression models’ feature coefficients.⁶ Coefficients with the largest absolute values are associated with more discriminating features.

We first discuss features from LIWC, our best performing feature set (Table 3a). Second-person pronouns (“2-you”) occurring around 20% and 30% into the negotiation are strongly predictive of a win for either party. An example use by the seller is “i think we can negotiate and talk about it but yeah you can start by giving me a price”. Similarly, it appears to be effective for buyers who bid their time and allow the seller to drive the negotiation by using filler words such as “mhm”, “k”, “yep”, and “huh” (LIWC category “netspeak”), especially near the beginning of the conversation. One interpretation could be that the buyer signals a comfort with silence and a lack of eagerness, effectively transferring the onus of driving the conversation onto the seller. This opens up the opportunity for the seller to consider the possibility of ceding ground on their asking price. For both buyer and seller, playing the ball into the other person’s court appears effective.

Furthermore, LIWC category “space” is associated with seller success. This category consists of seller-spoken words early in the negotiation like “small”, “location”, “big”, and “room”, among many others, which are used in reference to various aspects of the house. Early discussion of this sub-

⁶We use the average coefficients of the five models in cross validation.

10%	30%	50%	70%	90%
SELLER WIN				
b1-focuspast, s1-compare, b1-adj, s1-time, s2-social	s2-space, b3-posemo, s3-social, s2-you, s2-quant	b4-negemo, s2-space, s5-money, b1-focuspast, s2-you	b4-negemo, s2-space, b4-money, s2-you, s2-drives	s2-space, b4-negemo, s8-bio, s2-interrog, s8-body
BUYER WIN				
s1-adverb, s1-interrog, b1-i, b1-netspeak, b1-motion	b1-motion, b1-netspeak, b3-you, b3-bio, s1-interrog	b1-motion, s1-you, b3-bio, b1-netspeak, s1-interrog	s4-focuspast, s1-conj, b1-netspeak, b3-bio, s1-interrog	b3-bio, b1-netspeak, b1-motion, s1-interrog, s9-netspeak

(a) LIWC.

10%	30%	50%	70%	90%
SELLER WIN				
b-push, b-push b-new, b-repeat b-push, b-new, b-new b-push	b-push b-compare, b-new b-compare, b-push, b-compare b-repeat, b-push b-repeat	b-new b-compare, b-new b-repeat, b-push b-compare, b-push, b-push b-new	b-push b-compare, b-new b-compare, b-new b-allow, b-push b-allow, b-allow b-push	b-push b-compare, b-push b-new, b-repeat b-push, b-push, b-new b-compare
BUYER WIN				
b-new b-compare, b-repeat, b-push b-compare, b-compare b-push, b-compare	b-allow b-compare, b-allow, b-compare b-allow, b-compare b-push, b-compare	b-allow b-compare, b-new b-push, b-compare b-push, b-repeat b-new, b-new	b-compare, b-repeat b-allow, b-allow, b-new, b-allow b-compare	b-repeat b-allow, b-allow b-compare, b-compare b-allow, b-allow, b-compare b-compare

(b) Bargaining acts.

Table 3: Top features predicting negotiation success in each feature group, separated by role. Each column corresponds to the fraction of the conversation represented in the input. Prefixes “s-” and “b-” denote seller and buyer speech, respectively. The digit in a prefix refers to the features location within the negotiation (e.g. “b4” refers to buyer speech in first 40%). Top and bottom half of each table correspond to seller and buyer win features, respectively.

Buyer: Okay well I really like the house but I think that The price of \$235,000 is a bit excessive especially considering um the prices of some homes that are nearby The house I’m interested in that are selling for a lot less than that Um So I would definitely want to negotiate the price Um
Seller: Yeah How much how much was asking price again I believe it was 240
Buyer: Okay I think that a fair price would be around 218,000 Just considering other houses in the area
Seller: Um But like we also have like houses newly decorated we have like two fireplaces We also have a large eat in kitchen with all the appliances And uh comparing we all the house has uh 1,846 sq ft of space and which is more than the other first listing in appendix two

Table 4: Example transcript excerpt.

ject by the seller establishes a baseline justification for the housing price. This appears to play in favor of the seller as it clarifies the pretense that the asking price in question is substantive and sets the tone for a negotiation based on concrete considerations.

One of the more salient LIWC features for buyer is “b4-negemo”, predictive of a seller win. Prominent member words include “unfortunately”, “prob-

lem”, “sorry”, “lower”, and “risk”, suggesting that it is ineffective to frame the negotiation explicitly as a conflict resolution rather than a positive cooperative process. Given that the buyer requires movement on the asking price to succeed, they should avoid language that explicitly acknowledges that the seller may be compromising their interests. This result echoes the important role of negative expressions on negotiation outcomes by Barry (2008).

Another notable observation is that buyer-only bargaining acts are more predictive. To better make sense of this observation, Table 3b shows important features when predicting only with buyer bargaining act unigrams and bigrams. Most notably, *new offers* and *pushes* followed by *comparisons* consistently appear as two of the most influential features predictive of seller wins.

We present an example excerpt in Table 4 to illustrate such sequences. In this case, the comparison is serving the role of justifying the buyer’s new offer of \$218,000. This scenario often occurs the first time that a comparison is made by

either party: It provides the seller the opportunity to provide counter-evidence in favor of dismissing or devaluing the buyer’s offer. Tying an offer to a rationalization based on extrinsic evidence is ineffective for the buyer since its relative importance may be debated and new evidence may be introduced to weaken the effect of the buyer’s argument. This conclusion complements the finding that, in contrast to the buyer, the seller is advantaged when discussing details of the property, as evidenced by the LIWC feature “s2-space”.

Further Evaluation. As an additional experiment, we train a logistic regression model on the CRAIGSLISTBARGAIN dataset (He et al., 2018) and test it on our dataset. We include seller and buyer text, and use the same text encoding procedure described in §5.1. In the CRAIGSLISTBARGAIN dataset, the seller asking price is considered to be the seller’s private value for the item being sold and the buyer’s private value is separately specified. We consider the negotiation to be a seller win if the agreed price is higher than the midpoint between the two private values and a buyer win otherwise. Despite CRAIGSLISTBARGAIN having a significantly larger training dataset, the maximum test accuracy across all 10 fractions of our negotiations dataset is 54%, whereas we achieve a maximum of 60% accuracy when we train and test on our dataset. This experiment underscores the distinctiveness of our dataset and suggests that it may contain relevant linguistic differences to other datasets within the bargaining domain.

6 Conclusion

In this work we design and conduct a controlled experiment for studying the language of bargaining. We collect and annotate a dataset of *alternating offers* and *natural language* negotiations. Our dataset contains more turns per negotiation than existing datasets and, since participants communicate orally, our setting facilitates a more natural communication environment. Our dataset is further enhanced with annotated bargaining acts. Our statistical analyses and prediction experiments confirm existing findings and reveal new insights. Most notably, the ability to communicate using language facilitates cooperation and results in higher agreement and faster convergence. Both sellers and buyers benefit from maintaining a neutral role and providing the space for the other to formulate and justify new offers. Buyers should further avoid 1) using

house comparisons as primary rationales for their interests and 2) suggesting that by negotiating the asking price down, they are introducing conflict into the interaction.

Limitations

We note several important limitations of this work. Perhaps most importantly, our dataset is “naturalistic,” but not actually “natural” in the sense of independently occurring in the world. Though the interactions between our participants are real, the task itself is ultimately artificially constructed. In a real-world negotiation over something as valuable and significant as a house, the negotiating parties will be much more invested in the outcome than our experimental participants, whose actions change their outcome to the order of a few dollars. This difference in turn could lead real-world negotiating parties to speak differently and possibly employ substantially different strategies than we observe.

Methodologically, our study has a few important limitations. Firstly our analyses are based entirely on language that has been automatically transcribed (with some manual checks), and while this helps with expense and scale, these transcripts could be missing important subtleties that influence the outcome. Koenecke et al. (2020) uncover an important limitation of these systems, finding significant racial disparities in the quality of ASR transcriptions. The linguistic feature analysis we perform should be treated as largely exploratory, and provides suggestive and correlational rather than causal evidence for the relationship between language in the interactions and negotiation outcomes.

Lastly, there are further linguistic and interactional phenomena at play that we have not yet integrated into the analysis presented here. For one, we have access to the audio channel of participants’ actual speech, but we have not analyzed it in this work. There could very well be acoustic cues in participants’ speech that are as significant to the interactions as the textual features analyzed here, particularly speech prosody which has been shown to communicate social meanings that could be highly relevant to negotiation like friendliness (Jurafsky et al., 2009). This particularly extends to more interactional questions of not simply who said what, but what was said in response to what and in what way. For instance, existing research has shown that acoustic entrainment in dialog (e.g.,

interlocutor adaptation to one another in terms of prosody) has important social associations with dialogue success (Levitan et al., 2012). We leave a deeper investigation of these phenomena for future work.

Broader Impacts

This research, collectively with prior and future related work, has the potential to advance our understanding of negotiation, a ubiquitous human activity. Our dataset can enable future research into the dynamics of human bargaining as well as interpersonal interactions more broadly. By employing the findings and insights gained from such research, individuals may be able to enhance their ability to negotiate effectively in various settings, such as salary negotiations, personal relationships, and community initiatives. Meanwhile, we must acknowledge that while a better understanding of language as an instrument in social interaction can be empowering, it may also be used as a tool for manipulation.

Acknowledgements

We are grateful to Jessica Halten for helping us run the experiment through the Yale SOM Behavioral Lab. The experiment also would not have been possible without the excellent study session coordination by Sedzornam Bosson, Alexandra Jones, Emma Blue Kirby, Vivian Wang, Sherry Wu, and Wen Long Yang. We thank Rajat Bhatnagar for developing the web application used in the study. The human subjects experiment in this research was deemed exempt by the Yale University Institutional Review Board (IRB #2000029151). We thank Allison Macdonald and Sammy Mustafa for their effort in the data annotation process. Their work was an invaluable contribution to the success of this research project. We thank all members of the Chicago Human+AI Lab and LEAP Workshop for feedback on early versions of this work. Finally, we thank all anonymous reviewers for their insightful suggestions and comments.

References

Nicholas Asher, Julie Hunter, Mathieu Morey, Benamara Farah, and Stergos Afantenos. 2016. [Discourse structure and dialogue acts in multiparty dialogue: the STAC corpus](#). In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2721–2727, Portorož,

Slovenia. European Language Resources Association (ELRA).

Lawrence M. Ausubel, Peter Cramton, and Raymond J. Deneckere. 2002. [Chapter 50 bargaining with incomplete information](#). volume 3 of *Handbook of Game Theory with Economic Applications*, pages 1897–1945. Elsevier.

Bruce Barry. 2008. [Negotiator affect: The state of the art \(and the science\)](#). *Group decision and negotiation*, 17(1):97–105.

Max H. Bazerman, Jared R. Curhan, Don A. Moore, and Kathleen L. Valley. 2000. [Negotiation](#). *Annual Review of Psychology*, 51(1):279–314.

Iz Beltagy, Matthew E. Peters, and Arman Cohan. 2020. [Longformer: The long-document transformer](#). *CoRR*, abs/2004.05150.

Sally Blount. 2000. [Buying a house](#). *Dispute Resolution Research Center*.

Man Kit Chang and Carson C. Woo. 1994. [A speech-act-based negotiation protocol: Design, implementation, and test use](#). *ACM Trans. Inf. Syst.*, 12(4):360–382.

Vincent P Crawford. 1990. [Explicit communication and bargaining outcome](#). *American Economic Review*, 80(2):213–219.

Michael A. Fligner and Timothy J. Killeen. 1976. [Distribution-free two-sample tests for scale](#). *Journal of the American Statistical Association*, 71(353):210–213.

Yao Fu, Hao Peng, Tushar Khot, and Mirella Lapata. 2023. [Improving language model negotiation with self-play and in-context learning from ai feedback](#).

Andrew F. Hayes and Klaus Krippendorff. 2007. [Answering the call for a standard reliability measure for coding data](#). *Communication Methods and Measures*, 1(1):77–89.

He He, Derek Chen, Anusha Balakrishnan, and Percy Liang. 2018. [Decoupling strategy and generation in negotiation dialogues](#). In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 2333–2343, Brussels, Belgium. Association for Computational Linguistics.

Martha Jeong, Julia Minson, Michael Yeomans, and Francesca Gino. 2019. [Communicating with warmth in distributive negotiations is surprisingly counterproductive](#). *Management Science*, 65(12):5813–5837.

Dan Jurafsky, Rajesh Ranganath, and Dan McFarland. 2009. [Extracting social meaning: Identifying interactional style in spoken conversation](#). In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics, NAACL '09*, page 638–646, USA. Association for Computational Linguistics.

- Allison Koenecke, Andrew Nam, Emily Lake, Joe Nudell, Minnie Quartey, Zion Mengesha, Connor Touns, John R. Rickford, Dan Jurafsky, and Sharad Goel. 2020. [Racial disparities in automated speech recognition](#). *Proceedings of the National Academy of Sciences*, 117(14):7684–7689.
- Alice J Lee and Daniel R Ames. 2017. [“i can’t pay more” versus “it’s not worth more”](#): Divergent effects of constraint and disparagement rationales in negotiations. *Organizational Behavior and Human Decision Processes*, 141:16–28.
- Rivka Levitan, Agustín Gravano, Laura Willson, Štefan Beňuš, Julia Hirschberg, and Ani Nenkova. 2012. [Acoustic-prosodic entrainment and social behavior](#). In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 11–19, Montréal, Canada. Association for Computational Linguistics.
- Mike Lewis, Denis Yarats, Yann Dauphin, Devi Parikh, and Dhruv Batra. 2017. [Deal or no deal? end-to-end learning of negotiation dialogues](#). In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2443–2453, Copenhagen, Denmark. Association for Computational Linguistics.
- Rebecca Passonneau. 2006. [Measuring agreement on set-valued items \(MASI\) for semantic and pragmatic annotation](#). In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC’06)*, Genoa, Italy. European Language Resources Association (ELRA).
- Dean G Pruitt. 2013. *Negotiation behavior*. Academic Press.
- Alvin E Roth. 2020. [Bargaining experiments](#). In *The Handbook of Experimental Economics*, pages 253–348. Princeton University Press.
- Jeffrey Z Rubin and Bert R Brown. 1975. *The social psychology of bargaining and negotiation*. Elsevier.
- Roderick I. Swaab, William W. Maddux, and Marwan Sinaceur. 2011. [Early words that work: When and how virtual linguistic mimicry facilitates negotiation outcomes](#). *Journal of Experimental Social Psychology*, 47(3):616–621.
- Yla R. Tausczik and James W. Pennebaker. 2010. [The psychological meaning of words: Liwc and computerized text analysis methods](#). *Journal of Language and Social Psychology*, 29(1):24–54.
- Douglas P Twitchell, Matthew L Jensen, Douglas C Derrick, Judee K Burgoon, and Jay F Nunamaker. 2013. [Negotiation outcome classification using language features](#). *Group Decision and Negotiation*, 22(1):135–151.
- Hans Weigand, Mareike Schoop, Aldo de Moor, and Frank Dignum. 2003. [B2b negotiation support: The need for a communication perspective](#). *Group Decision and Negotiation*, 12(1):3–29.
- Haolan Zhan, Yufei Wang, Tao Feng, Yuncheng Hua, Suraj Sharma, Zhuang Li, Lizhen Qu, and Gholamreza Haffari. 2022. [Let’s negotiate! a survey of negotiation dialogue systems](#). Working paper. California State University, Northridge, CA.
- Tianyi Zhang, Felix Wu, Arzoo Katiyar, Kilian Q. Weinberger, and Yoav Artzi. 2021. [Revisiting few-sample bert fine-tuning](#).

Appendix

A Negotiation Excerpts

<p>Buyer: Okay well I really like the house but I think that The price of \$235,000 is a bit excessive especially considering um the prices of some homes that are nearby The house I'm interested in that are selling for a lot less than that Um So I would definitely want to negotiate the price Um</p> <p>Seller: Yeah How much how much Where the app was asking price again I believe it was 240</p> <p>Buyer: Okay I think that a fair price would be around 218,000 Just considering other houses in the area</p> <p>Seller: Um But like we also have like houses newly decorated we have like two fireplaces We also have a large eat in kitchen with all the appliances And uh comparing we all the house has uh 1,846 sq ft of space and which is more than the other first listing in appendix two</p>	<p>Buyer: My name is [name] Um I am an investor looking to buy a single household family in the neighborhood Um and your house based on the information that I was given seemed like a good option And I was looking at the housing market in the area and it seems like one of the houses that closely resembles your own house has been sold for \$213,000 Um so I am interested in buying your house at a price somewhere close to that Uh price</p> <p>Seller: Okay perfect Um Well um That house that you're talking about was actually sold quite a while ago so the prices have appreciated quite a bit and now the asking price that we have is \$240,000</p> <p>Buyer: Yeah</p>	<p>Buyer: I do feel like even though I agree it's a nice area it's a bit overpriced Um I mean speaking of comparisons the one I'm looking at right now listing 89 I was 6898 The selling price they're asking for is approximately 213,000 Um it has 1715 square feet And I've done the math That's a difference of 131 sq ft The difference in your asking price And my offering is to 27,000 So that equates to about \$206 per square foot Um That's the difference and I think that's a reasonable difference to make</p> <p>Seller: Yeah the market has been weirdly slow around here lately Um So we could come down slightly uh into the high two thirties let's say 239</p> <p>Buyer: Um I'll raise it 214</p> <p>Seller: Mhm Um Right we're gonna Stick with 239 I think</p>
---	---	---

Table 5: Push following by comparison examples

B Controlled Experiment

Compensation details summary. Each subject received \$10 for showing up and could earn additional bonus money per round. Bonus earnings depended on bargaining outcomes to incentivize subjects to engage in realistic negotiating behavior. Buyers could earn \$1 in bonus for every \$1,000 that the agreed sale price was *below* the buyer's private value of \$235,000, up to a maximum of \$10 in bonus money. Sellers could earn \$1 in bonus for every \$1,000 that the agreed sale price was *above* the seller's private value of \$225,000, up to a maximum of \$10. Given the private values of buyers and sellers, \$10 of surplus was available to split. No party earned bonus money in a round if an agreement was not reached.

C Logistic Regression Features

	Roles	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
LIWC	Buyer+Seller	266	296	409	547	687	824	962	1105	1244	1381
	Buyer	120	135	205	272	343	412	482	553	622	688
	Seller	146	161	204	275	344	412	480	552	622	693
Transcription Texts	Buyer+Seller	261	589	1052	1522	1979	2420	2385	2423	2397	2375
	Buyer	140	303	519	734	946	1161	1376	1554	1728	1869
	Seller	121	286	533	788	1033	1293	1493	1735	1916	2116
Bargaining Acts	Buyer+Seller	36	65	83	93	98	105	106	108	108	110
	Buyer	12	22	26	27	28	29	29	29	29	30
	Seller	14	23	29	32	33	33	33	33	33	33

Table 6: Logistic Regression Feature Counts

D Hyperparameters

Features	n-gram	Inner/Outer k -Folds	Max Iterations	ℓ_2 Coefficient
Numerical/BOW/LIWC	1	5	10k	$\{2^x x \in \{-10, -9, \dots, 0, 1\}\}$
Bargaing Acts	2	5	10k	$\{2^x x \in \{-10, -9, \dots, 0, 1\}\}$

Table 7: Logistic Regression hyperparameters. Unless otherwise specified, we use the default parameters from the Scikit-Learn LogisticRegression API.

Model	Speaker Role	k -Folds	Max Epochs	Batch Size	Optimizer	Learning Rate
LED	Seller + Buyer	5	20	4	AdamW	5e-5

Table 8: LongformerEncoderDecoder hyper-parameters. We used 3 epoch patience for early stopping based on training loss. We also implement best-practice recommendations from [Zhang et al. 2021](#) for few-sample BERT fine-tuning.


Table 9: Demographic Attributes of Study Subjects

	Percent (%)	Percent (%)
Gender		
Male	38.01	20.86
Female	60.23	9.55
Other	1.75	3.12
Age		
18-24	63.74	65.30
25-34	27.49	0.19
35-44	4.09	0.58
45-54	1.75	
55-64	2.34	
65-74	0.58	
Hispanic	13.45	
Race		
American Indian or Alaska Native	2.15	14.15
Asian	35.23	40.67
Black or African American	14.29	14.54
Native Hawaiian or Other Pacific Islander	0.20	13.16
White	48.73	10.22
Education		
< High School	0.00	0.00
High School or GED	12.09	1.36
Some college, no degree	36.26	7.41
Associate degree	2.14	14.62
Bachelor's degree	31.19	15.40
Master's degree	15.20	13.06
Doctorate or professional degree	3.12	15.79
Marital Status		
Single (never married)	84.80	15.98
Married or domestic partnership	13.45	9.55
Divorced	1.75	3.51
		3.31
		513
		521

Notes. This table reports select demographic attributes of study subjects. Attributes were collected from a survey of subjects prior to the start of each study session. Responses were voluntary. Participants were allowed to select multiple choices for Race. All other attribute questions allowed only a single choice response. Risk preferences were elicited from the question: "Are you generally a person who is willing to take risks or do you try to avoid taking risks?" Respondents rated themselves on a ten-point scale from 0 (unwilling to take risks) to 10 (very willing to take risks). The percentage of respondents in each demographic category is reported, except for the number of subjects, which are the raw counts of the number of participants in the experiment across all study sessions for whom we have demographic information and the number of experiment participants in total.


E Recruitment and instruction material

Table 9 reports select demographic attributes of study subjects.



Yale SOM Behavioral Research Lab

Online Study Available




Yale SOM researchers need participants for an online audio Zoom study. Participants must have an active Zoom account and a computer or laptop (no mobile devices please). This study should take 45-60 minutes to finish.

You will earn \$10-\$30 (depending on choices made within the study) as an Amazon eGiftcard. You must provide a valid email address and complete all components of the study in order to receive your eGiftcard payment, which will be sent within 2 business days after your participation.


Please note: There is an audio component to this study (no video required) and participants will be asked to upload an audio file upon completion of the study.

[Sign Up for the Housing Study - Virtual](#)

After signing up for your timeslot, you will have a secured spot for the study. The day prior to your session, you will receive a link for the study and a study user ID via the email address you have provided within SONA. This Zoom link will be equipped with a virtual waiting room and the RA will allow participants to enter the study at the session start time.

 Yale SCHOOL OF MANAGEMENT
Behavioral Lab

Yale School of Management Behavioral Lab
Copyright © 2020 Yale University • All rights reserved



HOUSING NEGOTIATION STUDY

WELCOME

Hello everyone, thank you for your patience as we waited for everyone to arrive. I am the study leader. You are about to participate in a study on negotiation, and you will be paid for your participation via an Amazon eGiftcard, privately emailed to you by the Yale SOM Behavioral Lab within two business days after the conclusion of the study.

Please close any program that you may have open on your computer besides Zoom. We will start with a brief instruction period.

If you have any questions during this period, please privately message the question to me, and I will answer it so that everyone can hear. In the chat, I will now send the weblink to the study instructions. Please follow along as I read the instructions.

GENERAL

In this study, you will negotiate the price to buy or sell a house with other participants. The study consists of two rounds of negotiation. The person you negotiate with in round 1 will differ from the person you negotiate with in round 2. In each round, one of you will play the role of the house buyer, whereas the other will play the role of the house seller. In both rounds, you will play the same role as either buyer or seller.

COMPENSATION

I will now describe the compensation. You will receive 10 dollars for participating in the study plus have the potential to earn up to 10 dollars bonus money in each round. The amount of bonus money you earn in each round depends on the outcome of the negotiation. Your total earnings for the study are the amount that you accumulate over the two rounds. The maximum cumulative earnings are 30 dollars, whereas the minimum cumulative earnings are 10 dollars.

CONSENT FORM + DEMOGRAPHIC SURVEY

I will now share a weblink to the consent form to participate in this study and a short demographics survey for you to complete. Please leave this Zoom session open while you complete the survey. The consent form will ask you to enter your study ID, so please be ready to enter it. If you do not consent to participating in

the study or if you are under age 18, please inform me via a private message. Once you complete the consent form and the demographics survey, please write a private message to me with the single word “done.”

ROLE PROMPTS

In a private message, I will now share information about the role of either the buyer or the seller to each of you. Take some time to read over this information, and use this information as you like in each negotiation round. As you read this information over, please keep this Zoom session open. Once you are finished reading the information, please send me a private message with the single word “done.”

ROUND 1

We will now begin the first round of the study. In this round, you will negotiate with another participant by only exchanging price offers and counter offers for the house. No other form of communication is allowed.

This negotiation will take place within a web application whose weblink I will share shortly. Upon clicking the weblink, you will begin exchanging offers with the other person. The buyer will propose the initial offer. If possible, please use the Google Chrome browser to open the weblink. While you negotiate, please keep this Zoom session open.

You and the other person will have a maximum of 15 minutes to negotiate, but you may finish before that time elapses. Once you complete the negotiation, privately message me the single word “done.” If you do not reach an agreement after 15 minutes, neither of you will earn bonus money for this round.

If the role you see in the weblink differs from the one I gave you earlier, please let me know.

Please wait to begin until after I say so once all the weblinks have been sent out.

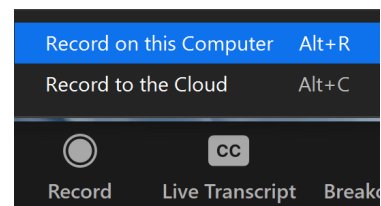
ROUND 2

Now we will begin the second round. In this round, you will negotiate with a different person. Pairs of participants will be assigned to individual breakout rooms to negotiate. In your breakout room, you will play the same role as either the buyer or

seller as you did in the first round. But now, you and the person you are paired with will negotiate the house price by talking to each other over Zoom audio only. The conversation is not limited to the exchange of price offers. Keep your video off the entire time. The buyer should begin the negotiation.

You and the other person will have up to 15 minutes to negotiate, but you may finish before that time elapses. If you do not reach an agreement within 15 minutes, neither of you will earn bonus money for this round.

Before you start negotiating, BOTH of you should RECORD the breakout room session. To record, click record in the meeting controls at the bottom of your screen. Record to your Computer.



Please DO NOT start recording until after you have entered the breakout room. You may start negotiating once you hit RECORD. Once you finish negotiating, STOP the recording. Once you stop recording, leave the breakout room and return to the main room.



END OF ROUND 2

The second round of negotiation is complete. Please click the weblink to a survey I will send shortly. This survey will give instructions to upload your audio recording, if you consent to do so.

Please keep this Zoom session open while you complete the survey. Once you finish the survey, privately message me the single word "done."

Before uploading your recorded audio file to the survey, please rename it "studyID.mp4" without the quotation marks, where studyID is your Study ID.

Survey to upload audio recording:

https://yalesurvey.ca1.qualtrics.com/jfe/form/SV_5opHhEOIRAigful

Role of Buyer

(Based on "Buying a House" by Sally Blount,
Northwestern Kellogg Dispute Resolution Research Center)

The text on this page is available only to the buyer.

Housing values have risen rapidly in Centerville over the last few years, and you are interested in investing in a piece of real estate. Optimally, you would like to find a single family home in the \$220,000 to \$235,000 range, which you could rent for a few years and then resell at a profit. You recently saw an advertisement in the "Centerville Review" for a house near Centennial Park (see Appendix 1 below), which is being sold directly by its owner. Based upon the description, this house seemed like the type of investment that you are seeking. You arranged to visit the home last week. The asking price was \$240,000 and you were favorably impressed.

You have since collected information on comparable houses to help you assess the worth of this house (see Appendix 2 below). You have decided that you would like to buy the house, but not at a price in excess of \$235,000. In fact, you would like to buy the house at a price as close to \$220,000 as possible. However, you would be willing to pay up to \$235,000 before walking away from this opportunity. You will meet with the owner today to discuss buying the house.

You cannot share the following information about your compensation with the seller. If the study coordinators learn that you have shared the following information in any form, you will forfeit your compensation.

If you and the owner reach an agreement, you will earn \$1 in bonus for every \$1,000 that the agreed sale price is below your walk-away price of \$235,000, up to a maximum of \$10 in bonus money.

You will not earn any bonus money if you do not reach an agreement with the owner or if you agree to a price above \$235,000.

Appendix 1

(available to both buyer and seller)

Single House
Listing # 90 13878

Description

- 4 bedrooms + 1 recreation room + 2.5 bathrooms
- Split-level style
- Built in 1947
- 1846 square feet of space

Inside Amenities

- Finished hardwood floors
- 2 fireplaces
- Master bedroom with an entire wall of closets plus master bath
- Large eat-in kitchen with all appliances
- Newly decorated

Outside Amenities

- Comfortable & updated brick
- Beautiful landscaping
- Fenced backyard and mature trees
- Detached garage (for 2.5 cars)
- Restaurants and transportation within walking distance
- Near Hastings & Centennial parks

Asking Price: \$240,000

Appendix 2

(available to both buyer and seller)

Prices of neighboring homes with similar characteristics

Listing #	Selling Price	Square Feet
89 06898	\$213,300	1715
89 04725	\$233,600	1875
89 08614	\$239,600	1920

Role of Seller

(Based on “Buying a House” by Sally Blount,
Northwestern Kellogg Dispute Resolution Research Center)

The text on this page is available only to the seller.

You have owned your house near Centennial Park in Centerville for several years (see Appendix 1 below). You originally purchased it for \$155,000. To save on commissions, you have decided to sell the house yourself. After discussions with your friends who are real estate investors, you have set an asking price of \$240,000. The house has been on the market for one month, and you have not yet had a firm offer.

You have always believed that you have one of the nicest houses in the Centennial Park area. You also think your house is favorably priced in comparison to comparable homes in Centerville. It has been several months since the last house was sold in the Centennial Park area. Thus, your asking price on a per square foot basis is higher (see Appendix 2 below).

Since your house has been on the market for several weeks, you have decided that you would settle for any offer that yielded at least \$225,000. However, you would prefer to sell as close to \$240,000 as possible. You would rather hold on to the house than sell below \$225,000. Last week a prospective buyer visited your home and showed a keen interest in buying the house. You will meet with that prospective buyer today to discuss selling the house.

You cannot share the following information about your compensation with the buyer. If the study coordinators learn that you have shared the following information in any form, you will forfeit your compensation.

If you and the buyer reach an agreement, you will earn \$1 in bonus for every \$1,000 that the agreed sale price is above your minimum sale price of \$225,000, up to a maximum of \$10 in bonus money.

You will not earn any bonus money if you do not reach an agreement with the buyer or if you agree to a price below \$225,000.

Appendix 1

(available to both buyer and seller)

Single House
Listing # 90 13878

Description

- 4 bedrooms + 1 recreation room + 2.5 bathrooms
- Split-level style
- Built in 1947
- 1846 square feet of space

Inside Amenities

- Finished hardwood floors
- 2 fireplaces
- Master bedroom with an entire wall of closets plus master bath
- Large eat-in kitchen with all appliances
- Newly decorated

Outside Amenities

- Comfortable & updated brick
- Beautiful landscaping
- Fenced backyard and mature trees
- Detached garage (for 2.5 cars)
- Restaurants and transportation within walking distance
- Near Hastings & Centennial parks

Asking Price: \$240,000

Appendix 2

(available to both buyer and seller)

Prices of neighboring homes with similar characteristics

Listing #	Selling Price	Square Feet
89 06898	\$213,300	1715
89 04725	\$233,600	1875
89 08614	\$239,600	1920

Welcome to the negotiation bidding page! You are the buyer. The negotiation process starts after you propose a price.

Role of Buyer

(Based on "Buying a House" by Sally Blount, Northwestern Kellogg Dispute Resolution Research Center)

Housing values have risen rapidly in Centerville over the last few years, and you are interested in investing in a piece of real estate. Optimally, you would like to find a single family home in the \$220,000 to \$235,000 range, which you could rent for a few years and then resell at a profit. You recently saw an advertisement in the "Centerville Review" for a house near Centennial Park (see Appendix 1), which is being sold directly by its owner. Based upon the description, this house seemed like the type of investment that you are seeking. You arranged to visit the home last week. The asking price was \$240,000 and you were favorably impressed.

You have since collected information on comparable houses to help you assess the worth of this house (see Appendix 2). You have decided that you would like to buy the house, but not at a price in excess of \$235,000. In fact, you would like to buy the house at a price as close to \$220,000 as possible. However, you would be willing to pay up to \$235,000 before walking away from this opportunity. You will meet with the owner today to discuss buying the house.

You cannot share the following information with the seller. If the study coordinators learn that you have shared the following information in any form, you will forfeit your compensation.

If you and the owner reach an agreement, you will earn \$1 in bonus for every \$1,000 that the agreed sale price is below your walk-away price of \$235,000, up to a maximum of \$10 in bonus money.

You will not earn any bonus money if you do not reach an agreement with the owner or if you agree to a price above \$235,000.

Time Remaining: 15:00

Figure 6: Buyer web app page.

Welcome to the negotiation bidding page! You are the seller. The negotiation process starts after the buyer proposes a price.

Role of Seller

(Based on "Buying a House" by Sally Blount, Northwestern Kellogg Dispute Resolution Research Center)

You have owned your house near Centennial Park in Centerville for several years (see Appendix 1). You originally purchased it for \$155,000. To save on commissions, you have decided to sell the house yourself. After discussions with your friends who are real estate investors, you have set an asking price of \$240,000. The house has been on the market for one month, and you have not yet had a firm offer.

You have always believed that you have one of the nicest houses in the Centennial Park area. You also think your house is favorably priced in comparison to comparable homes in Centerville. It has been several months since the last house was sold in the Centennial Park area. Thus, your asking price on a per square foot basis is higher (see Appendix 2).

Since your house has been on the market for several weeks, you have decided that you would settle for any offer that yielded at least \$225,000. However, you would prefer to sell as close to \$240,000 as possible. You would rather hold on to the house than sell below \$225,000. Last week a prospective buyer visited your home and showed a keen interest in buying the house. You will meet with that prospective buyer today to discuss selling the house.

You cannot share the following information with the buyer. If the study coordinators learn that you have shared the following information in any form, you will forfeit your compensation.

If you and the buyer reach an agreement, you will earn \$1 in bonus for every \$1,000 that the agreed sale price is above your minimum sale price of \$225,000, up to a maximum of \$10 in bonus money.

You will not earn any bonus money if you do not reach an agreement with the buyer or if you agree to a price below \$225,000.

Time Remaining: 15:00

Waiting for other person's response

Figure 7: Seller web app page.

ACL 2023 Responsible NLP Checklist

A For every submission:

- A1. Did you describe the limitations of your work?
7
- A2. Did you discuss any potential risks of your work?
7
- A3. Do the abstract and introduction summarize the paper’s main claims?
Abstract and 1
- A4. Have you used AI writing assistants when working on this paper?
Left blank.

B Did you use or create scientific artifacts?

3, 4, 5

- B1. Did you cite the creators of artifacts you used?
5
- B2. Did you discuss the license or terms for use and / or distribution of any artifacts?
1
- B3. Did you discuss if your use of existing artifact(s) was consistent with their intended use, provided that it was specified? For the artifacts you create, do you specify intended use and whether that is compatible with the original access conditions (in particular, derivatives of data accessed for research purposes should not be used outside of research contexts)?
We provide a link to a webpage for requesting access to the artifacts we create. This form includes information on the intended use.
- B4. Did you discuss the steps taken to check whether the data that was collected / used contains any information that names or uniquely identifies individual people or offensive content, and the steps taken to protect / anonymize it?
At time of publication, we are in the process of preparing the data for restricted release via form. We will detail additional information as appropriate in the form for requesting access.
- B5. Did you provide documentation of the artifacts, e.g., coverage of domains, languages, and linguistic phenomena, demographic groups represented, etc.?
1, 2, 7
- B6. Did you report relevant statistics like the number of examples, details of train / test / dev splits, etc. for the data that you used / created? Even for commonly-used benchmark datasets, include the number of examples in train / validation / test splits, as these provide necessary context for a reader to understand experimental results. For example, small differences in accuracy on large test sets may be significant, while on small test sets they may not be.
Left blank.

C Did you run computational experiments?

5

- C1. Did you report the number of parameters in the models used, the total computational budget (e.g., GPU hours), and computing infrastructure used?
We report the GPU used in section 5. We cite the model we used for experimentation, additional details may be found in the original paper introducing the LED model.

The Responsible NLP Checklist used at ACL 2023 is adopted from NAACL 2022, with the addition of a [question on AI writing assistance](#).

- C2. Did you discuss the experimental setup, including hyperparameter search and best-found hyperparameter values?
5 and Appendix D
- C3. Did you report descriptive statistics about your results (e.g., error bars around results, summary statistics from sets of experiments), and is it transparent whether you are reporting the max, mean, etc. or just a single run?
3,4,5
- C4. If you used existing packages (e.g., for preprocessing, for normalization, or for evaluation), did you report the implementation, model, and parameter settings used (e.g., NLTK, Spacy, ROUGE, etc.)?
We don't use any existing package for significant processing or evaluation. We do use huggingface tokenizer and cite the model we use.
- D** **Did you use human annotators (e.g., crowdworkers) or research with human participants?**
3
- D1. Did you report the full text of instructions given to participants, including e.g., screenshots, disclaimers of any risks to participants or annotators, etc.?
Appendix E
- D2. Did you report information about how you recruited (e.g., crowdsourcing platform, students) and paid participants, and discuss if such payment is adequate given the participants' demographic (e.g., country of residence)?
3
- D3. Did you discuss whether and how consent was obtained from people whose data you're using/curating? For example, if you collected data via crowdsourcing, did your instructions to crowdworkers explain how the data would be used?
3
- D4. Was the data collection protocol approved (or determined exempt) by an ethics review board?
3
- D5. Did you report the basic demographic and geographic characteristics of the annotator population that is the source of the data?
3